A SECURE WEB-BASED DIGITAL IMAGE LIBRARY

Mary Thompson, Srilekha Mudumbai, William Johnston

MRThompson@lbl.gov

http://imglib.lbl.gov

ERNEST ORLANDO LAWRENCE
BERKELEY NATIONAL LABORATORY

General Image Library problem

Many large images

- hundreds to thousands
- 1 - 50 MBytes
- Associated textual metadata and derived images

Images kept in centralized Mass Storage Systems or off-line storage

- on tape or removable disk- slow access
- remote from user - slow network
- access methods awkward and restrictive

Images generated at remote sites

- medical images - cardio-angiography, MRI, etc.
- electron microscopes

Image users at remote sites

- collaborators at other sites

Our Solution - A Web Server

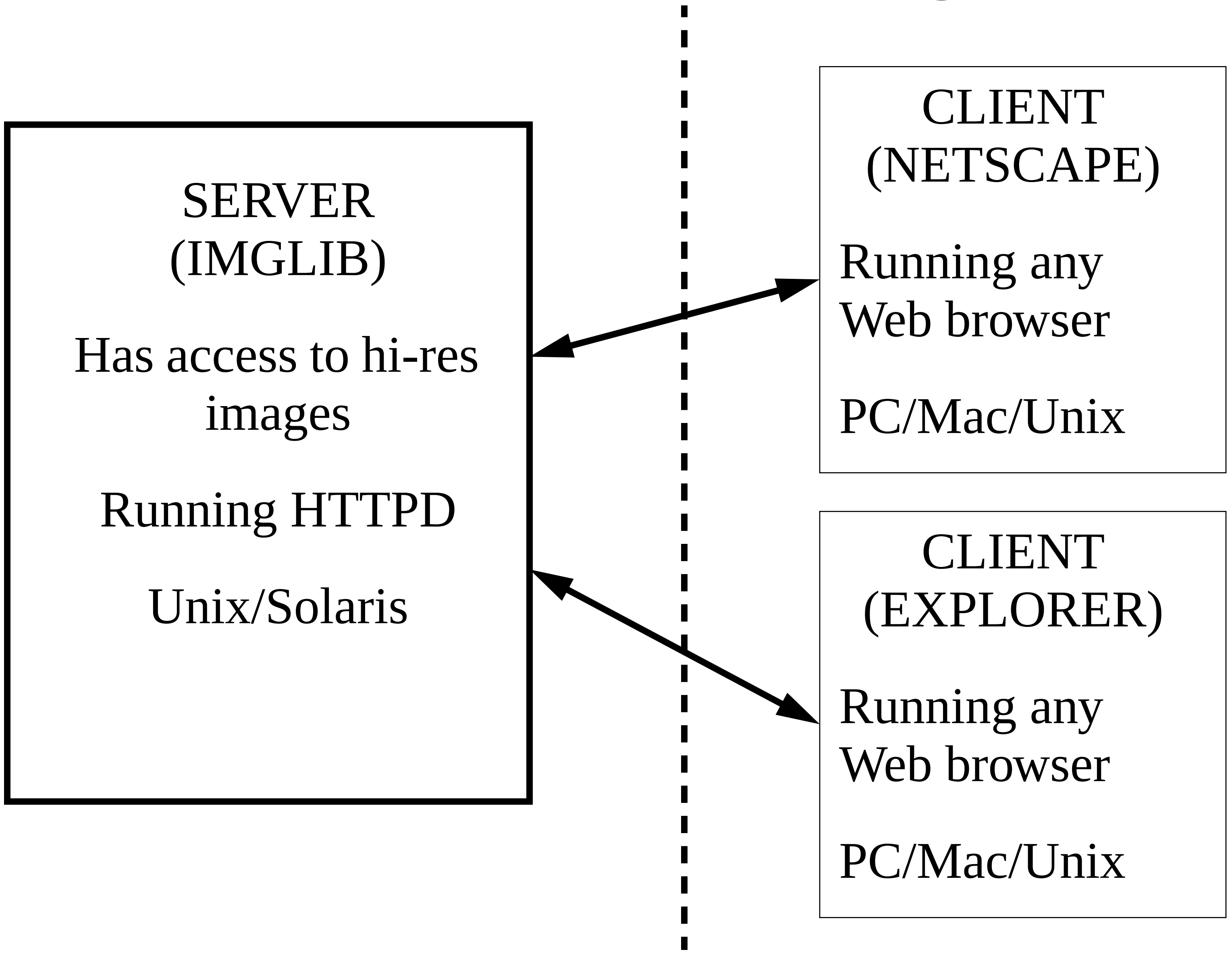
Solves two problems easily:

- user remote from the high-resolution images
- users at sites remote from each other sharing images

Server code is cgi scripts run by standard Web Server, e.g. NCSA HTTPD 4.1, Apache

Client code is standard Web browser, Netscape, MicroSoft Explorer

Client/Server paradigm



ImgLib Functional Overview

Cataloguing system

Creates and manages Metadata

- creates and displays derived thumbnail images
- indexes and searches text information

Provides interface to tertiary storage

Does not modify original images

ImgLib Functions

Smaller copy images are created and kept on-line on server machine

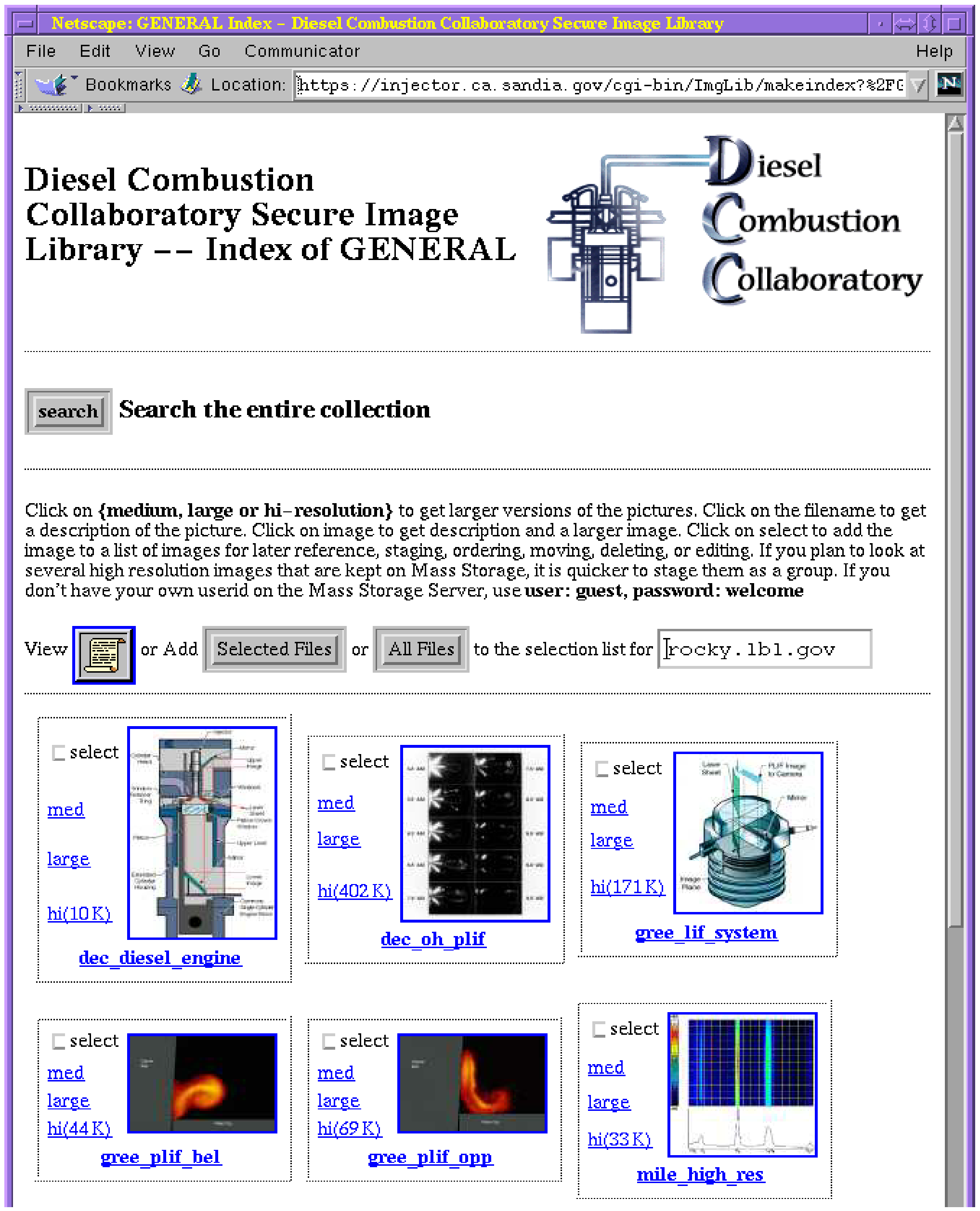
Collections are divided into logical hierarchical subcollections that are small enough to browse quickly with a Web browser

Provides a Web interface to get images from tertiary storage

Textual information about each image is stored and indexed to allow searching

All user or curator interactions are through Web

Browsing Interface


Netscape: GENERAL Index – Diesel Combustion Collaboratory Secure Image Library
File Edit View Go Communicator Help
Bookmarks Location: https://injector.ca.sandia.gov/cgi-bin/ImgLib/makeindex?%2F0

Diesel Combustion Collaboratory Secure Image Library –– Index of GENERAL

Diesel
Combustion
Collaboratory

search **Search the entire collection**

Click on **{medium, large or hi–resolution}** to get larger versions of the pictures. Click on the filename to get a description of the picture. Click on image to get description and a larger image. Click on select to add the image to a list of images for later reference, staging, ordering, moving, deleting, or editing. If you plan to look at several high resolution images that are kept on Mass Storage, it is quicker to stage them as a group. If you don't have your own userid on the Mass Storage Server, use **user: guest, password: welcome**

View or Add Selected Files or All Files to the selection list for rocky.lbl.gov

☐ select med large hi(10 K) **dec_diesel_engine**

☐ select med large hi(402 K) **dec_oh_plif**

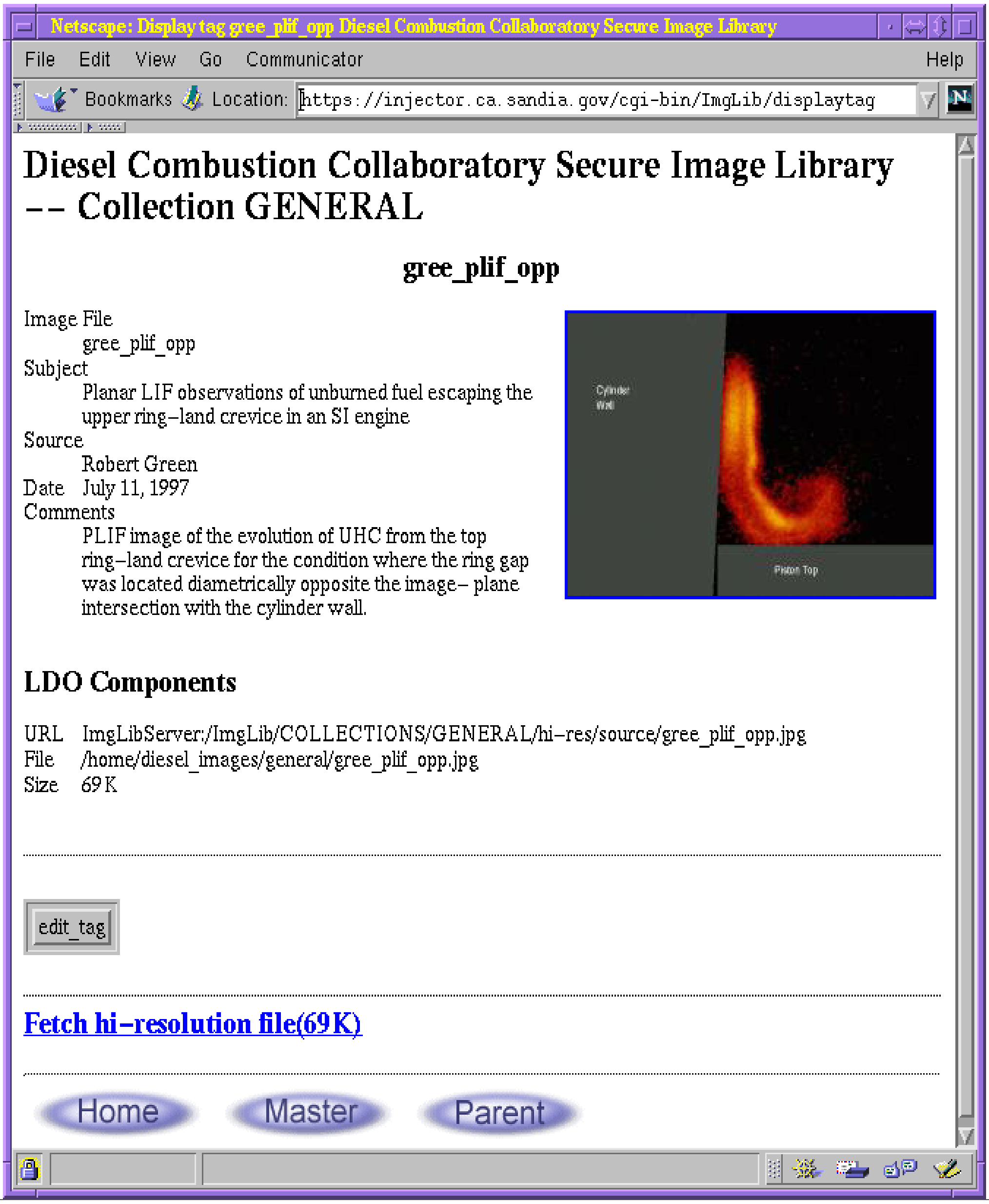
☐ select med large hi(171 K) **gree_lif_system**

☐ select med large hi(44 K) **gree_plif_bel**

☐ select med large hi(69 K) **gree_plif_opp**

☐ select med large hi(33 K) **mile_high_res**

Data Object Definition Document



Netscape: Display tag gree_plif_opp Diesel Combustion Collaboratory Secure Image Library
File Edit View Go Communicator Help
Bookmarks Location: https://injector.ca.sandia.gov/cgi-bin/ImgLib/displaytag

Diesel Combustion Collaboratory Secure Image Library -- Collection GENERAL

gree_plif_opp

Image File	gree_plif_opp
Subject	Planar LIF observations of unburned fuel escaping the upper ring–land crevice in an SI engine
Source	Robert Green
Date	July 11, 1997
Comments	PLIF image of the evolution of UHC from the top ring–land crevice for the condition where the ring gap was located diametrically opposite the image– plane intersection with the cylinder wall.

LDO Components

URL	ImgLibServer:/ImgLib/COLLECTIONS/GENERAL/hi-res/source/gree_plif_opp.jpg
File	/home/diesel_images/general/gree_plif_opp.jpg
Size	69 K

Fetch hi–resolution file(69 K)

Status

Lung microscopy collection

- 1400 1-4 Megabyte scanning electron microscope images
- the primary researcher and his group members use ImgLib routinely to find and view his images

Lab Photo Archive

- 1000 images on-line selected from the thousands of photos in the Berkeley Lab archive
- they have been used for a variety of Web publications and for general browsing by people in and outside the Lab

Kaiser Permanente angiograms

- data collected daily for all patients from Oakland and Vallejo primary care centers (June ‘96 thru Jan‘97)
- 2 weeks worth of data on a high-speed network memory cache (30 Gbytes) - 60 patient cardio-angiography studies consisting of about 800 individual video clips stored
- 3 months of data on tertiary storage
- doctors at primary care institutions have used these images and are evaluating the project

Original Implementation

NCSA HTTPD server

HTTPD cgi-scripts - Perl and C

- display dynamic index pages for a collection.
- display data-object definition documents
- fetch original image
- curating the collection

Implementation (cont)

Public library packages

- glimpse search engine from Univ. of Arizona - http://glimpse.cs.arizona.edu:1994/
- pbmplus image package, Jeff Poskanzer - http://www.acme.com/software/
- cjpeg conversion - Independent JPEG group - ftp from ftp.uu.net/graphics/jpeg

Access Control Limitations

Web servers control GET and POST actions

Use local .htaccess files with user/password or domain access control.

Passwords are passed over the net uuencoded (not encrypted)

My scripts mimics this method, added additional .htaccess files to control write actions and access to original images.

Requirements for Secure ImgLib

Data passed over the network must be encrypted

Users wanting access must be securely identified

Needs to work in a distributed collaborative environment

- Distributed stakeholders
- Distributed users
- Distributed access control

Want finer control over actions than just read and write

Encrypted Network Connections

Provided by Apache with SSL patches

- http//ww.apache.org
- ftp://ftp.psy.uq.oz.au/pub/Crypto/SSL
- ftp://ftp.ox.ac.uk/pub/crypto/SSL/ftp

Or buy a Netscape Web server

- FastTrack for $295.

Use a server private key and certificate to encrypt the connections

Client needs to do nothing except reference https

Securely Identify Users
(Provided by Netscape)

Require users to get an Identity Certificate from a trusted CA

- Netscape CA $525

Set server to require client authorization and tell it what CA's to trust.

Clients must now use Netscape Navigator and present their Identity Certificate with their requests. Navigator makes this straightforward.

Internet Explorer may use a different and incompatible method of handling Certificates.

Distributed Access Control

Distributed stakeholders create UseCondition certificates locally.

We provide a Java gui application that communicates with the resource server to help the stakeholder create a valid UseCondition and Attribute Certificates.

Certificates can be stored locally by the stakeholder but accessible via a Web or LDAP server or can be entered by the stakeholder into a central LDAP or Web Server.

Image Lib Resources are Hierarchical

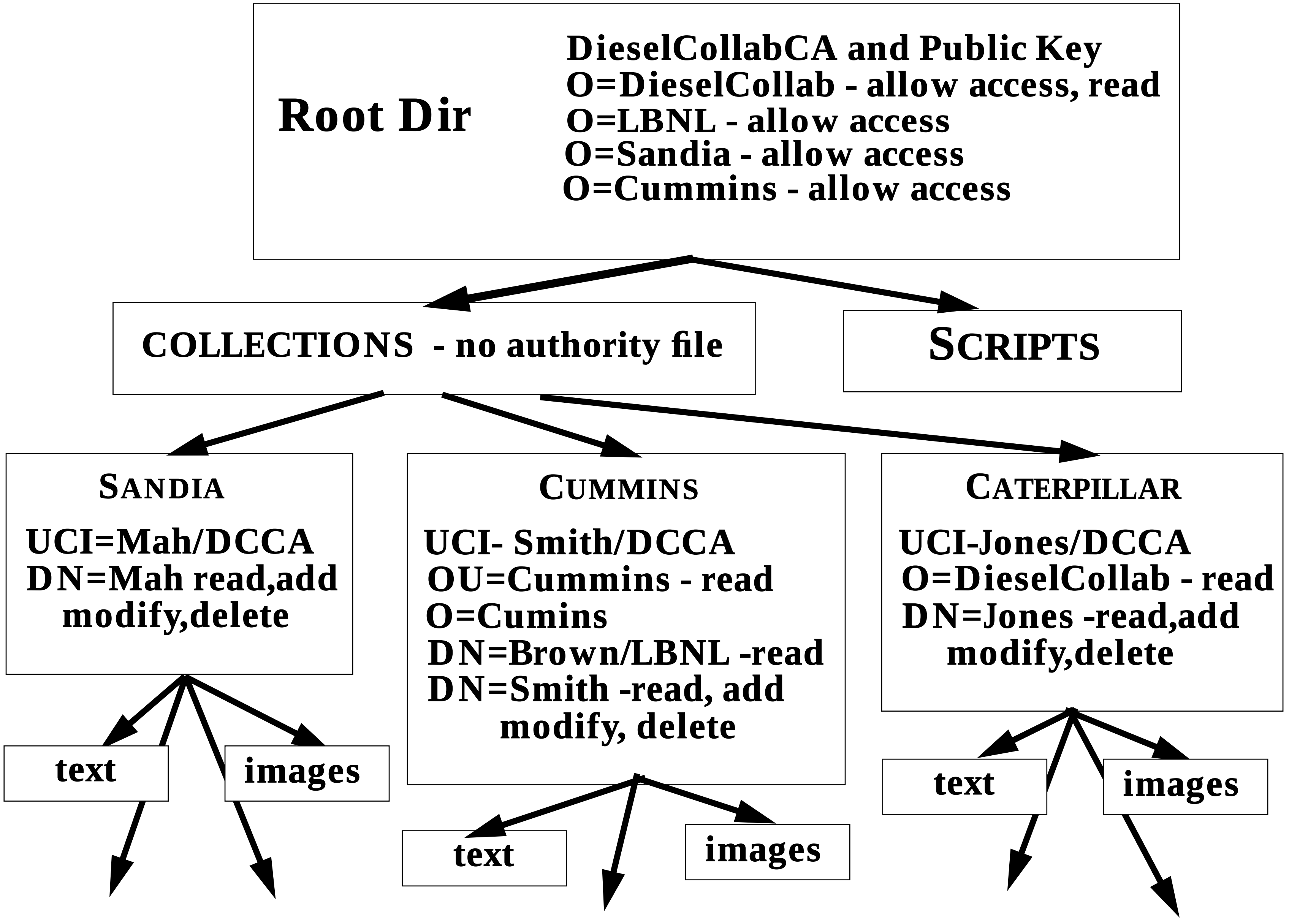
Collections are organized like directory trees.

Access control and curation tends to follow the same hierarchy.

The protection granularity for ImgLib is directories (aka collections of images).

We wish to be able to inherit access control from a parent directory as well as set it locally.

ImgLib Resource Architecture



Authority Files

Trusted starting point of access control

Files on the Server machine that can be associated with each resource to be protected

Consists of:

- places to look for certificates
- list of acceptable UseCondition issuers
- list of places to look for UseConditions

Root Authority File

Must exist in the root of the resource tree

Contains the list of CA's to trust and their ID certificates (public key)

Places to find UseConditions

At least one UseCondition that grants the broadest allowed access to the resource tree. Otherwise we fail-safe and no-one is allowed access to the tree.

ImgLib UseConditions

Name of the resource

Scope - local or sub-tree

Access and actions - read, add, delete, modify, fetch-hi-res

Conditional expression of attributes and values

Who can issue attribute certificates and their CA’s

CA’s to trust to verify the user identity

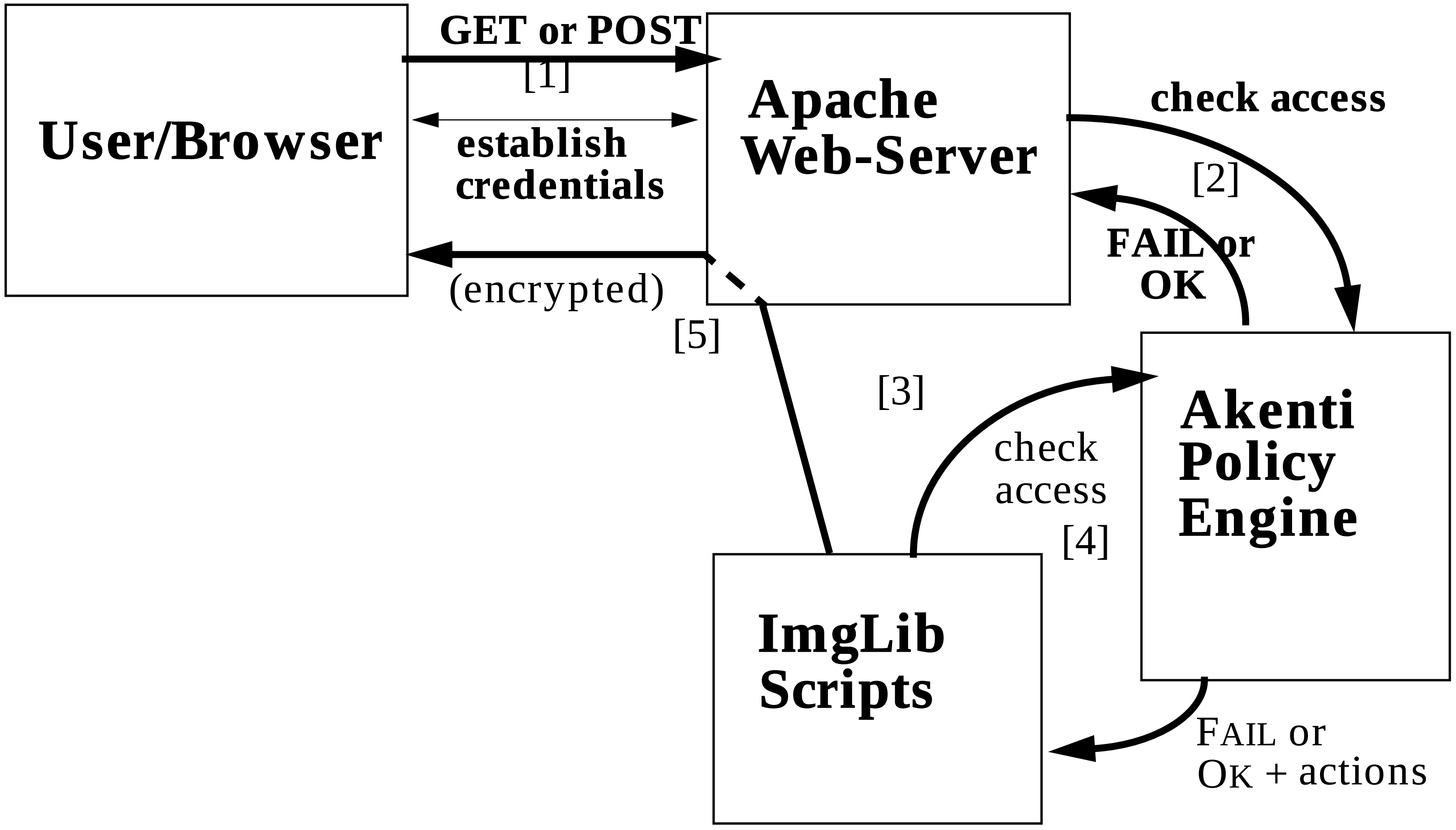
Finding all the UseConditions

Given a request for a resource, Akenti has to know what Authority Files apply

We wish to be able to inherit access from a parent directory as well as set it locally

- Search resource directory for .htauthority file
- Check each parent directory for an .htauthority file.
- Use all the relevant UseConditions that are found.

Making an access



Akenti Policy Engine steps

Akenti Check Access interface called with:

- **DN -** /C=US/O=DieselCollab/OU=LBNL/CN=Mary Thompson
- **RootDir -** /opt/akenti/akenti-htdocs
- **resourceDir** - COLLECTIONS/SANDIA
- **authFileName** - .htauthority
- **protocolHost** - https//www.diesel-colab.sandia.gov
- **allowedActions -** returned (access,read,add,delete,modify,fetch-hi-res)

Policy Engine steps (2)

Policy Engine starts at resourceDir looking for Authority Files and works up to RootDirectory.

For each AuthFile it encounters:

- gathers up all the UseConditions
  - Does LDAP lookups
  - pulls through URLs
  - may search hash directories returned by URL.

Policy Engine steps (3)

Builds a list of all relevant UseConditions

- Checks to see if the resource name in the UseCondition matches the resource we are referencing.
- Check that the scope includes the resource

Policy Engine steps (4)

For each UseCondition that enables access:

- Checks the issuer and signature for validity
- If the user fails to satisfies any such UseCondition, he is denied access to the resource.

To see what actions are allowed:

- starts from the bottom up and finds the first directory that contained UseConditions.
- Checks all those UseConditions for validity.
- See which ones the user satisfies.
- Combines all the rights granted by these UseConditions.

Status

Just starting deployment as part of the Diesel Combustion Collaboratory

To Do list

- Caching
  - Very short-term caching of permissions to a resource
  - Longer term caching of UseCondition and Attriubute certificates

To Do (cont)

- Vulnerabilty to network outage or insecure servers.
  - How to tell if UseConditions are not found
  - Use redundent distributed servers, and/or a reliable central server
  - Fail-safe or fail-soft?
- Allowing Authority Files to be managed remotely